Working Paper – May 27, 2016

Optimal Pricing under Mixed Logit Customer Choice

Ruben van de Geer[1]
Sandjai Bhulai
VU University Amsterdam, Faculty of Sciences, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

Abstract

In this paper we consider the problem of pricing multiple differentiated products. This is challenging as a price change in one product, not only changes the demand of that particular product, but also the demand for the other products. To address this problem, customer choice models have recently been introduced as these are capable of describing customer choice behavior across differentiated products. In the present paper the objective is to obtain the revenue-maximizing prices when the customer's decision making process is modelled according to a particular customer choice model, namely the mixed logit model. The main advantage of using the mixed logit model, also known as the random coefficients logit model, for this purpose is its flexibility. In the single-product case we establish log-concavity of the optimization problem under certain regularity conditions. In addition, in the multi-product case, we present the results of our extensive numerical experiments. These suggest that the mixed logit model, by taking unobserved customer heterogeneity and flexible substitution patterns into account, can significantly improve the attainable revenue.

1 Introduction

The problem of pricing a range of differentiated products is very common from a business perspective, but simultaneously very challenging, since a price change in one product, not only changes the demand of that particular product, but also the demand for the other products (so called, substitution). In an attempt to give substance to this problem, customer choice models have recently been introduced in a pricing framework, as these models have proven useful to describe customer choice behavior across differentiated products (e.g., Li and Huh, 2011; Gallego and Wang, 2014; Li et al., 2015; Huh and Li, 2015). These customer choice models, also known as discrete choice models, assume that customers assign utility to products based on the products' attributes and, subsequently, maximize their utility and purchase accordingly (or, possibly, choose not to purchase after all). This framework is probabilistic in a sense that customer choice models provide us with a probability distribution over the alternatives, thereby recognizing that, besides the observable attributes, the attained utility is also the consequence of random unobservable factors. The notion that customers value product attributes differently and choose accordingly seems particularly interesting when customers are offered a menu of differentiated products and expected revenue needs to be optimized with respect to prices for the full range of products. Besides its theoretical relevance, this seems practically appealing given the increased availability of customer-level purchasing data.

1.1 Research objective

The focus of the present paper is on maximizing the expected revenue with respect to prices in a multi-product setting when the customer's decision making process is modelled according to a particular customer choice model, namely the mixed logit model. The main advantage of using the mixed logit model, also known as the random coefficients logit model, for this purpose is its flexibility. In fact, it has been shown that the mixed logit model can approximate any customer choice model up to any desired level of precision (McFadden et al., 2000). With the development of simulation-based estimation methods, such as Markov chain Monte Carlo, the mixed logit model has become practically useful. Ever since, it has been extensively used in, e.g., marketing (e.g., Rossi et al., 1996) and transportation sciences (e.g., Hensher and Greene, 2003) to empirically analyze the decision making behavior of customers. From an optimization perspective, the main challenge is due to the revenue function, which includes a

[1]Corresponding author - r.vande.geer@vu.nl

multidimensional integral for which no closed-form solution exists, making the mixed logit model "not conducive for analysis" in a pricing framework (Akcay et al., 2010). In an attempt to overcome this, we rely on Laplace integration to obtain a closed-form approximation of the optimization problem.

1.2 Related literature

Existing work on price optimization under customer choice models seems to have been initiated by Hanson and Martin (1996), who showed that the expected revenue function under the multinomial logit (MNL) model is not concave with respect to the product prices. It has been shown, however, by Song and Xue (2007) and Dong et al. (2009) that the profit optimization problem under MNL is jointly concave with respect to the purchase probabilities,[2] which are injective with respect to the prices. Therefore, the optimal purchase probabilities can be obtained by solving a reformulation of the original optimization problem and then the optimal prices can be retrieved accordingly. More recently, Aydin and Porteus (2008) and Akcay et al. (2010) showed that the profit function under MNL with constant price sensitivities across products is unimodal with respect to the markups, i.e., the prices minus costs, so that solving the first-order conditions with respect to the markups is sufficient. Moreover, they show that the optimal markups are constant across all products. This implies that prices are constant across products under MNL if the cost is uniform (e.g., in case of revenue optimization). However, in an attempt to explain price variations across differentiated products often observed in practice, Akcay et al. (2010) show in a setting with dynamic prices and inventory constraints that the optimal markups vary across products and are higher than the static markups. Besides pricing under MNL, pricing under the nested logit (NL) model, a generalization of MNL, has received attention recently. The NL model assumes that customers first choose a *nest* (i.e., a subset) of products and, subsequently, choose a product from that particular nest. In Li and Huh (2011) and Gallego and Wang (2014) pricing under the NL model is studied, whereas Li et al. (2015) and Huh and Li (2015) consider a more general version of the model with an arbitrary number of levels of nests. Finally, only recently, Alptekinoğlu and Semple (2016) introduced the exponomial model for price and assortment optimization. This model has, in contrast to MNL and NL, a negatively skewed distribution of consumer utilities, which, in particular cases, describes the choice behavior of customers better than the MNL and NL specifications.

The aforementioned references reveal that the prevailing customer choice models used in a pricing framework are the MNL and the NL models. The former is arguably the most famous customer choice model and has received tremendous attention amongst both scientists and practitioners from diverse fields since it was first introduced by McFadden (1973). Despite its popularity, it suffers from three well-documented drawbacks, which can be resolved by the mixed logit model (Train, 2009). First of all, MNL assumes that if customers have identical observable characteristics (e.g., age or income), then they have identical attribute valuations. In other words, it is assumed that heterogeneity can only originate from observables and not from unobservable effects, such as latent personal traits. Second of all, MNL assumes that the unobserved part of a customer's attained utility is independent over time. This implies that when panel data is used for estimation, which is often the case in practice, serial dependence in choice behavior can only be induced by including lagged variables and, e.g., someones persisent latent preference for a certain unobservable cannot be captured. Finally, the last drawback pertains to what are known as substitution patterns. These patterns describe how the probability distribution over alternatives changes given a change in the composition of alternatives or the attributes of the alternatives. The MNL model assumes a very specific substitution pattern, known under the independence of irrelevant alternatives (IIA) property. This property implies that the relative probability of choosing one alternative over another, i.e., the ratio of these two probabilities, is independent of the composition of alternatives or the attributes of the alternatives. In our particular case, this would imply that an increase in the price

[2]These purchase probabilities are sometimes referred to as *market shares*, which are, apart from the interpretation, equivalent.

of one product, would increase the purchase probabilities of the other products with the same proportion (this is often referred to as *proportional* substitution). The NL model generalizes MNL and relaxes the IIA assumption that underlies MNL. In doing so, it is assumed that the alternatives can be partitioned into nests, so that the IIA property holds within these nests, but not across nests (McFadden et al., 1978; McFadden, 1980). However, the first two of the aforementioned drawbacks are not addressed by NL.

1.3 Contribution

This work's contribution lies in the introduction of the mixed logit model in a multi-product pricing framework. More precisely, we assume that the products' perceived quality is multivariate normally distributed, thereby introducing the notion of customer heterogeneity, as well as allowing for a wide variety of substitution patterns. Since the corresponding optimization problem comprises a multidimensional Gaussian integral for which no closed-form solution exists, we approximate the optimization problem by means of Laplace integration. This allows us to obtain the following structural and numerical results on the approximated optimization problem.

For the single-product case we show that the approximated optimization problem is in general not log-concave. However, we show that the approximated optimization problem is strictly log-concave when the variance is bounded by a quantifiable threshold. In addition, in a numerical study we focus on two aspects that are particularly interesting when considering the mixed logit model, namely the effect of unobserved customer heterogeneity and various substitution patterns on optimal prices and attainable revenue.

This paper is organized as follows. In Section 2 we study the problem of optimizing revenue with respect to prices under mixed logit choice and derive an approximation to this problem, for which we show it is log-concave in the single-product case. Then, in Section 3 we provide some of the results of our numerical study. Finally, in Section 4 we conclude by summarizing and providing guidance for future research.

2 Price Optimization under Mixed Logit Choice

2.1 Problem description

Let us formalize the pricing problem as follows. Assume there is a set $\mathcal{J}$ of available products (excluding the no-choice option) and let $J := |\mathcal{J}|$ be the cardinality of this set. We assume that customers assign utility to each product $j \in \mathcal{J}$ according to the following linear specification:

$$U_j = \alpha_j - p_j \beta_j + \varepsilon,$$

where ε is randomly i.i.d. standardized Gumbel distributed, p_j is the prevailing price, β_j is the price sensitivity, and α_j is a measure for the perceived quality of product j (which may, e.g., be a function of the product's attributes). For convenience, we will assume that $\beta_j = \beta$ throughout. Let us denote by U_0 the 'outside option', i.e., the utility that customers enjoy from not purchasing anything. By convention and without loss of generality we set $U_0 = \varepsilon$ (or, equivalently, $\alpha_0 = 0$ and $p_0 = 0$).

Under the MNL model, the parameters $\alpha := (\alpha_1, \alpha_2, \ldots, \alpha_J)$ and β are assumed to be fixed. The resulting choice probability for product $j \in \mathcal{J}$ has the following convenient closed-form solution, arguably being the primary reason for its popularity,

$$\begin{aligned} q_j(\alpha, p) :&= \mathbb{P}\left\{ \max_{k \in \mathcal{J}} U_k = U_j \right\} \\ &= \frac{e^{\alpha_j - p_j \beta}}{1 + \sum_{k \in \mathcal{J}} e^{\alpha_k - p_k \beta}} \end{aligned} \tag{1}$$

$$= \left[e^{-(\alpha_j - p_j \beta)} + \sum_{k \in \mathcal{J}} e^{\alpha_k - \alpha_j - (p_k - p_j)\beta} \right]^{-1},$$

where $p = (p_1, p_2, \ldots, p_J)$ and the explicit dependence on α is imposed to keep notation in line with further model development. In addition, the no-purchase probability equals $q_0(\alpha, p) = \left[1 + \sum_{k \in \mathcal{J}} e^{\alpha_k - p_k \beta}\right]^{-1}$.

The mixed logit model generalizes MNL by assuming that the model parameters are randomly distributed over the population of customers. In this particular study, we consider the case in which α is multivariate normally distributed with mean μ and covariance matrix Σ. We assume that β, the price sensitivity, is a fixed (non-random) coefficient. This implies that we incorporate customer heterogeneity with respect to the overall measure for quality of a product, whereas we assume that the customers are homogeneous with respect to price sensitivity.[3] All in all, the probability that a customer purchases product $j \in \mathcal{J} \cup \{0\}$ equals

$$s_j(p) := \int \left[e^{-(\alpha_j - p_j \beta)} + \sum_{k \in \mathcal{J}} e^{\alpha_k - \alpha_j - (p_k - p_j)\beta} \right]^{-1} (2\pi)^{-J/2} |\Sigma|^{-1/2} e^{-\frac{1}{2}(\alpha - \mu)^\top \Sigma^{-1} (\alpha - \mu)} d\alpha, \tag{2}$$

for which no closed-form solution is available. The corresponding optimization problem is

$$\pi^* := \max_p \pi(p), \tag{3}$$

where $\pi(p) = \sum_{j \in \mathcal{J}} s_j(p) p_j$. Therefore, the objective is to maximize revenue, which, without much additional effort, can be turned into a profit-maximizing optimization problem by subtracting marginal cost from the prices.

2.2 Motivating example

To motivate the use of a mixed logit model in a pricing framework, we consider a single-product toy example with $\alpha \sim \mathcal{N}(\mu = 1.0, \sigma^2 = 1.0)$ and $\beta = 0.1$. Then, under MNL or NL we would implicitly assume that α has a degenerate distribution with point mass at μ. The corresponding revenue-maximizing price under MNL equals $(W(e^{\mu-1})+1)/\beta = 15.67$ (e.g., Li and Huh, 2011), where W denotes the Lambert W function. To gain insight on the optimal price under mixed logit, we Monte Carlo-integrated the expected revenue $\pi(p)$ for a range of prices, of which the result is shown in Figure 1. The figure reveals that, at least in this particular case, unobserved heterogeneity with respect to the product's valuation causes the MNL price to be too low. Therefore, in an attempt to address such and similar problems, we proceed by discussing a closed-form approximation to (3). Getting ahead of ourselves, under this closed-form approximation we would find a price of 17.99 in the aforementioned example, which would yield a significant improvement.

2.3 Laplace integrating s_j

In an attempt to gain insight in optimal pricing under the mixed logit model, we rely on an approximation to s_j based on Laplace's method of integration and refer to this approximation by LA (Laplace approximation). LA in the current context is closely related to the work of Harding and Hausman (2007), who used a similar approximation for a different purpose, namely estimation of the mixed logit parameters, whereas we assume that the parameters are given and optimize revenue with respect to price. We begin

[3] The framework that will be discussed here allows for β to be normally distributed as well, however, this would imply that with positive probability β is negative, which has undesirable consequences when we are optimizing with respect to price. Namely, this would imply that the price of one (or more) of the products can be set to infinity to collect infinite revenues almost surely.

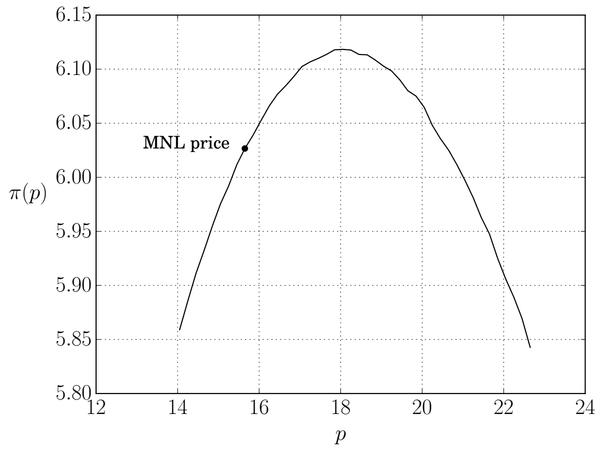

Figure 1: The expected revenue $\pi(p)$ when $\alpha \sim \mathcal{N}(\mu = 1.0, \sigma^2 = 1.0)$ and $\beta = 0.1$.

by rewriting (2) as follows:

$$s_j(p) = (2\pi)^{-J/2}|\Sigma|^{-1/2} \int e^{-\lambda g_j(\alpha,p)} d\alpha \tag{4}$$

where $\lambda = 1$ and

$$g_j(\alpha, p) = \frac{1}{2}(\alpha - \mu)^\top \Sigma^{-1}(\alpha - \mu) + \overbrace{\log\left(e^{-(\alpha_j - p_j\beta)} + \sum_{k\in\mathcal{J}} e^{\alpha_k - \alpha_j - (p_k - p_j)\beta}\right)}^{:=h_j(\alpha,p)}.$$

In the sequel, we will write $g_j(\alpha) = g_j(\alpha, p)$ and $h_j(\alpha) = h_j(\alpha, p)$ where possible to reduce the notational burden. It has been shown by Harding and Hausman (2007) that $g_j(\alpha)$ is convex in α, so that a global minimizer $\alpha^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_J^*)$ in $\mathbb{R}^J$ exists (the dependence of α^* on p and j is suppressed here). Assume for now that we know α^*, then the general idea of Laplace's method of integration is that when λ is 'large', the integral in (4) is dominated by a small neighbourhood around α^*. Therefore, we proceed by Taylor-expanding $g_j(\alpha)$ quadratically around its minimum α^*. For that purpose, let

$$g_j'(x) := \left.\frac{\partial g_j(\alpha)}{\partial\alpha}\right|_{\alpha=x} \quad \text{and} \quad g''(x) := \left.\frac{\partial^2 g_j(\alpha)}{\partial\alpha\partial\alpha^\top}\right|_{\alpha=x}$$

i.e., the gradient and Hessian of g_j, respectively, evaluated at x (the latter is independent of j, as will become clear below). Then,

$$\begin{aligned} g_j(\alpha) &\approx g_j(\alpha^*) + (\alpha - \alpha^*)^\top g_j'(\alpha^*) + \frac{1}{2}(\alpha - \alpha^*)^\top g''(\alpha^*)(\alpha - \alpha^*) \\ &= g_j(\alpha^*) + \frac{1}{2}(\alpha - \alpha^*)^\top g''(\alpha^*)(\alpha - \alpha^*), \end{aligned} \tag{5}$$

since $g_j'(\alpha^*) = 0$ as α^* is a minimum. When we plug (5) into (4) we obtain

$$\begin{aligned} \widetilde{s}_j(p) :&= |\Sigma|^{-1/2} e^{-\lambda g_j(\alpha^*)} \int (2\pi)^{-J/2} e^{-\frac{\lambda}{2}(\alpha-\alpha^*)^\top g''(\alpha^*)(\alpha-\alpha^*)} d\alpha \\ &= |\Sigma|^{-1/2} |\lambda g''(\alpha^*)|^{-1/2} e^{-\lambda g_j(\alpha^*)} \underbrace{\int (2\pi)^{-J/2} |\lambda g''(\alpha^*)|^{1/2} e^{-\frac{\lambda}{2}(\alpha-\alpha^*)^\top g''(\alpha^*)(\alpha-\alpha^*)} d\alpha}_{=1} \end{aligned}$$

$$= |\lambda A|^{-1/2} e^{-\frac{\lambda}{2}(\alpha^*-\mu)^\top \Sigma^{-1}(\alpha^*-\mu)} \left[e^{-(\alpha_j^*-p_j\beta)} + \sum_{k\in\mathcal{J}} e^{\alpha_k^*-\alpha_j^*-(p_k-p_j)\beta} \right]^{-1} \tag{6}$$

where

$$A = I_J + \Sigma\left(\operatorname{diag}(\vec{q}(p)) - \vec{q}(p)\vec{q}(p)^\top\right)$$
$$\vec{q}(p) := (q_1(p), \ldots, q_J(p))^\top,$$
$$q_j(p) := q_j(\alpha^*, p).$$

By construction, this approximation does not guarantee that $\sum_{j\in\mathcal{J}\cup\{0\}} \widetilde{s}_j(p) = 1$, i.e., that it is a proper probability mass function. Therefore, the approximations should be scaled accordingly.

All in all, we have obtained a closed-form approximation $\widetilde{s}_j(p)$ for the J-dimensional integral presented in (2). For this approximation it can be shown that

$$s_j(p) = \widetilde{s}_j(p) \cdot (1 + O(\lambda^{-1})),$$

which implies that for $\lambda \to \infty$, it holds that $\widetilde{s}_j \to s_j$. Since in our particular case $\lambda = 1$, the relative error is $O(1)$, i.e., the relative error is bounded by a constant, which is not of much practical use. However, it has been observed that the approximation is reasonably accurate in a *sub-asymptotic* setting, i.e., when λ is low (Butler et al., 2002; Harding and Hausman, 2007; Asmussen et al., 2014). Inspired by these works together with the notion that alternatives, such as Monte Carlo integration, seem not appropriate for optimization purposes, we proceed with LA and assess its performance numerically.

It still remains to obtain $\alpha^* \in \mathbb{R}^J$, i.e., the global minimizer of g_j. Since it can be shown that $g_j(\alpha)$ is convex, it suffices to solve $g_j'(\alpha) = 0$ for α, i.e., α^* is the solution to

$$\Sigma^{-1}(\alpha - \mu) + \frac{\partial h_j(\alpha)}{\partial \alpha} = 0 \tag{7}$$

where

$$\frac{\partial h_j(\alpha)}{\partial \alpha} = \Big(q_1(\alpha, p), q_2(\alpha, p), \ldots, q_j(\alpha, p) - 1, \ldots, q_J(\alpha, p)\Big)^\top.$$

This system of equations cannot be solved for α analytically. Therefore, we will rely on available efficient numerical root-finding algorithms to obtain α^*.[4]

2.4 Price optimization under LA

Let us consider the optimization problem when approximating the mixed logit probabilities with LA. In doing so, define the total revenue under price vector p as $\widetilde{\pi}(p) := \sum_{j\in\mathcal{J}} \widetilde{s}_j(p)p_j$, then,

$$\widetilde{\pi}^* := \max_p \widetilde{\pi}(p) \tag{8}$$

represents the maximal attainable revenue. In the single-product case, the optimization problem in (8) is in general not log-concave since, for example, when $\alpha \sim \mathcal{N}(\mu = 20, \sigma^2 = 400)$ and $\beta = 5.0$, then $0.626 = \log\widetilde{\pi}(3.75) < \frac{1}{2}\log\widetilde{\pi}(3.25) + \frac{1}{2}\log\widetilde{\pi}(4.25) = 0.638$. However, under some conditions it can be shown that (8) is strictly log-concave, as is shown in the following proposition.

Proposition 1. *In the single-product case with $\alpha \sim \mathcal{N}(\mu, \sigma^2)$ the optimization problem in* (8) *is strictly log-concave for $\sigma^2 \leq 34.92$.*

[4]Alternatively, (7) can be Taylor-approximated to obtain an approximated closed-form solution (Harding and Hausman, 2007). However, then, in general $g_j'(\alpha^*) \neq 0$ in (5).

Proof. Let α^* be the solution to $\sigma^2 = (\alpha - \mu)\left(1 + e^{\alpha - p\beta}\right)$ and let $q(p) = q(\alpha, p)|_{\alpha=\alpha^*}$ and $q_0(p) = q_0(\alpha, p)|_{\alpha=\alpha^*}$, then, after some algebra (for details see Appendix A) it can be derived that

$$\frac{\partial \log \widetilde{\pi}(p)}{\partial p} = \frac{1}{p} - \beta \frac{\overbrace{\left(\frac{\sigma^2}{2} + \alpha^* - \mu\right) q_0(p)q(p) + \sigma^2(\alpha^* - \mu)[q_0(p)q(p)]^2 + q_0(p)}^{=f}}{\underbrace{[1 + \sigma^2 q_0(p)q(p)]^2}_{=g}}. \tag{9}$$

We proceed by showing that (9) is decreasing in p by showing that $\frac{\partial f}{\partial p} g - f \frac{\partial g}{\partial p} > 0$ under the hypothesis that $\sigma^2 \leq 34.92$. It can be shown that

$$\frac{\partial f}{\partial p} g - f \frac{\partial g}{\partial p} > 0$$
$$\Leftrightarrow$$
$$q(p) + 3q(p)^2 - 8q(p)^3 + 6q(p)^5 - 2q(p)^6 + \frac{1}{\sigma^2} + \frac{2}{\sigma^4} > 0$$
$$\Leftrightarrow$$
$$\sigma^2 \leq 34.92$$

where the last step results from Sturm's theorem, which allows us to show that when $\sigma^2 = 34.92$ the LHS has no roots for $q(p) \in [0, 1]$. The claim then follows from the observation that LHS increases when σ^2 decreases. □

3 Numerical Study

3.1 Single-product price optimization

In the following section, we consider single-product price optimization under the mixed logit choice model. More precisely, we numerically assess the cost of assuming that the population of customers is homogeneous, i.e., that α is fixed at value μ, while, in fact, the population of customers is heterogeneous, i.e., α is normally distributed with expectation μ and variance σ^2. In doing so, given parameter values β, μ, and, σ^2, we maximize revenue with respect to price under customer homogeneity by using MNL and under customer heterogeneity by using LA. The optimal price under MNL is equal to

$$\widetilde{p}_{MNL} = \frac{W(e^{\mu-1}) + 1}{\beta} \quad \text{(e.g., Li and Huh (2011)).}$$

In addition, let $\widetilde{p}_{LA}$ be the optimal price under LA obtained by numerically optimizing the logarithm of (8), which is a strictly concave maximization problem according to Proposition 1. Thereafter, we approximate the *true* revenue under $\widetilde{p}_{MNL}$ and $\widetilde{p}_{LA}$, i.e., $\pi(\widetilde{p}_{MNL})$ and $\pi(\widetilde{p}_{LA})$, by Monte Carlo integrating (2) with appropriate convergence conditions. For $\beta = 0.5$, $\mu = 1.0$, and varying values of σ^2 the revenue, price, and purchase probability are plotted in Figure 2.

Several interesting comments regarding Figure 2 can be made. First of all, for very low variance ($\sigma^2 \approx 0$) the optimal price under MNL and LA approximately coincide, as one would expect as the normal distribution degenerates as $\sigma^2 \to 0$. In addition, since the price under MNL is independent of σ^2, we observe a constant price under MNL at $\widetilde{p}_{MNL} \approx 3.13$, which is strictly smaller than $\widetilde{p}_{LA}$. The main insight, however, follows from the observation that under LA, the optimal price increases as the variance increases, i.e., when the customers become increasingly heterogeneous. This increase in optimal price under LA for increasing variance results in an increase in revenue both absolute and relative to MNL. Presumably, when variance is relatively high, by increasing the price one can increase the revenue from

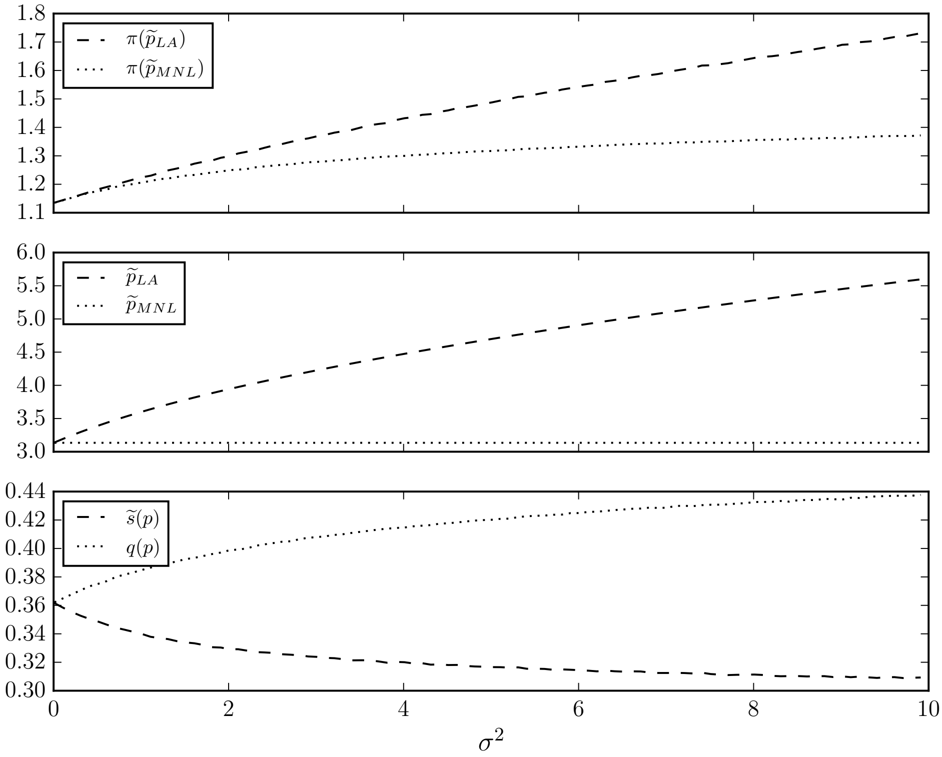

Figure 2: The top, middle, and bottom plot represent the revenue, price, and purchase probability, respectively, under LA and MNL for $\beta = 0.5$, $\mu = 1.0$, and various values of σ^2. The dashed graphs pertain to LA and the dotted graphs pertain to MNL.

customers that value the product relatively high, whilst this increase outweighs the lost revenue from customers with a relatively low valuation of the product. However, it seems unjustified to conclude that, in general, the revenue-maximizing price —which remains unknown— increases in variance or is strictly larger than $\widetilde{p}_{MNL}$, since properties of the exact optimization problem in (3) remain undiscovered in this work.

3.2 Multi-product price optimization

In the previous section, the implied cost of ignoring customer heterogeneity in the single-product case was analyzed. In the current section, we numerically assess in a multi-product setting how various model specifications, i.e., various settings for μ and Σ, affect revenue. The multi-product setting is especially worth considering, since optimal prices under MNL are all equal when optimizing with respect to revenue (e.g., Akcay et al. (2010); Li and Huh (2011)). In other words, in the two-dimensional case, under MNL it holds that $p_1 = p_2$, even if the products are differentiated in a sense that their perceived quality differs, i.e., that $\alpha_1 \neq \alpha_2$. This seems hard to justify from the business perspective, where an assortment of differentiated products is in general not priced uniformly.

For illustrative purposes, let us consider the case of two products with parameter settings $\beta = 0.2$, $\mu = (2, 4)^\top$, and

$$\Sigma_c = \begin{pmatrix} 5 & \mathrm{c} \\ \mathrm{c} & 5 \end{pmatrix}. \tag{10}$$

As such, the parameter c regulates the dependency between perceived quality of the two products. In particular, it seems interesting to consider the two border cases, i.e., $c = -4.99$ and $c = 4.99$, which

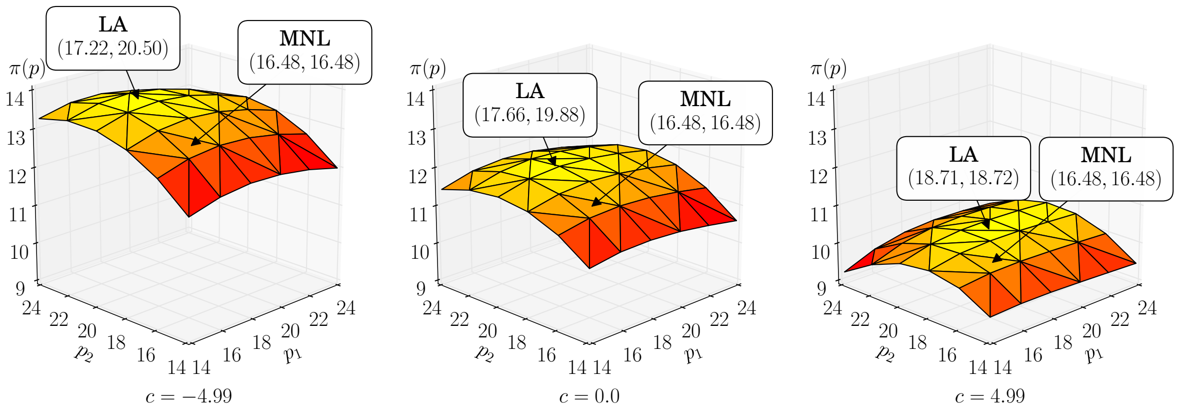

Figure 3: Revenue as a function of price with $\beta = 0.2$, $\mu = (2, 4)^\top$, and Σ_c for $c \in \{-4.99, 0, 4.99\}$. In the figure, the optimal prices under LA and MNL are indicated. The revenue $\pi(p)$ is obtained by means of Monte Carlo integration.

represent, respectively, (near) perfect negative and positive correlation between α_1 and α_2. For these cases, together with $c = 0$, i.e., independent α_1 and α_2, the revenue functions with corresponding optimal prices are illustrated in Figure 3. This figure reveals the following notable aspects. First of all, it can be observed that —at least in this particular case— the maximal revenue under LA is larger than under MNL and that the maximal revenue decreases when c increases from -4.99 to 0 and, subsequently, to 4.99. Presumably, the decrease of potential revenue in c may be attributed to the fact that with increasing c, the differentiation amongst products decreases, so that the product assortment becomes less attractive overall. In other words, loosely speaking, when c is 'low', most customers 'like' either one of the products, whereas, when c is 'high', some customers like both products and some dislike both products, leading to lower (potential) revenue. This also leads us to belief that the framework presented here has interesting implications for assortment optimization under customer choice models, which is a related optimization problem that received considerable attention in the recent decades (see, e.g., Rusmevichientong et al. (2014) for assortment optimization under mixed logit with the parameters being randomly distributed with finite support).

Second of all, when considering LA, as c increases we observe that the optimal prices converge to a single price, i.e., for $c \nearrow 5.00$ it holds that $p_1 \approx p_2$. This can be observed from the third plot in Figure 3, where $c = 4.99$ and $(p_1, p_2) = (18.71, 18.72)$, as well as from Figure 4 where the relationship between p_1 and p_2 and c under LA is illustrated. This occurs not only in this particular case, but occurred for all parameter configurations that were considered during extensive numerical experiments. Presumably, since the bivariate normal distribution degenerates as $c \nearrow 5.00$, prices coincide in a similar fashion as under the MNL model.

4 Conclusion

In the current work, we introduced the mixed logit model for the purpose of pricing multiple differentiated products. Until now, attention has primarily been on two particular choice models, namely MNL and NL, for which it has proven to be challenging to develop structural results, such as (log-)concavity of the profit or revenue function and characterizations of optimal prices for various model specifications. The mixed logit model can —in contrast to aforementioned models— account for unobserved customer heterogeneity and allows for flexible substitution patterns. However, the model is notoriously hard to analyze, as the purchase probabilities, and thus the revenue function, comprises a multidimensional Gaussian integral. In an attempt to address this challenge we approximate the problem by means of

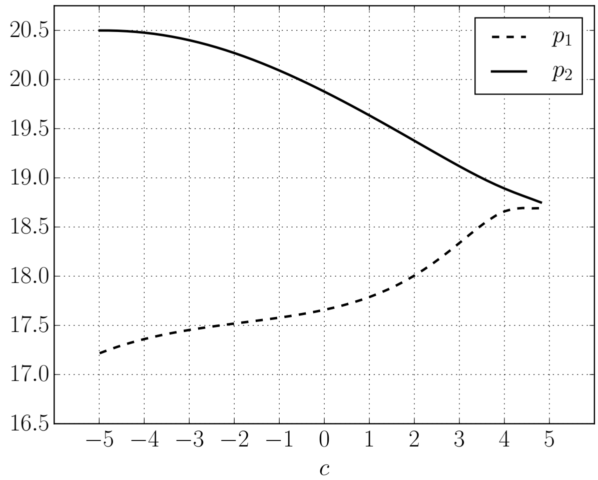

Figure 4: Optimal prices under LA as c increases with $\beta = 0.2$, $\mu = (2, 4)^\top$, and Σ_c as in (10).

Laplace integration.

In the single product case, we were able to establish log-concavity under certain distributional conditions, more specifically on the variance of the parameter that measures the perceived quality of the product. The extensive work required to establish this, suggests that proving a similar claim in the multi-product case is not feasible, at least with the current approach. The multi-product case, however, has been analyzed numerically. More specifically, for illustrative purposes we considered the two-product instance and considered the effect of unobserved customer heterogeneity and various substitution specifications on the optimal prices and attainable revenue. In general, the numerical experiments indicate that increased customer heterogeneity with respect to perceived quality leads to higher attainable revenue. Presumably, extra revenue can be earned from the customers that value the product(s) relatively high. Similarly, increased negative covariance between the perceived quality of two products also increases the attainable revenue, which may be contributed to the increased overall attractiveness of the assortment.

Although it has proven to be burdensome to analyze the mixed logit model in a pricing framework, we believe that the model may be appreciated for its flexibility and, therefore, its capability to reflect realistic phenomena. For future research, many questions remain unanswered, e.g., regarding (log-)concavity of the revenue function with multiple products and the effect of product-specific price sensitivities. In addition, our work may encourage others to introduce the mixed logit model in an assortment optimization framework, where the decision variable is not price, but on selecting the optimal subset of products to offer on a certain occasion.

Appendix A: Proof of Proposition 1

Preliminaries
Some useful expressions are the following:

$$\frac{\partial \alpha}{\partial p} = \frac{(\alpha-\mu)e^{\alpha-p\beta}\beta}{1+(\alpha-\mu+1)e^{\alpha-p\beta}} \qquad \text{(by implicit differentiation)}$$

$$\frac{\partial \alpha}{\partial p} - \beta = -\beta\frac{1+e^{\alpha-p\beta}}{1+(\alpha-\mu+1)e^{\alpha-p\beta}}$$

$$\frac{\partial q_0(\alpha,p)}{\partial p} = \frac{q(\alpha,p)\beta}{1+(\alpha-\mu+1)e^{\alpha-p\beta}}$$

$$\frac{\partial q(\alpha,p)}{\partial p} = -\frac{\partial q_0(\alpha,p)}{\partial p}$$

$$\frac{\partial[q_0(\alpha,p)q(\alpha,p)]}{\partial p} = \frac{\beta q_0(\alpha,p)q(\alpha,p)(1-2q_0(\alpha,p))}{1+\sigma^2 q_0(\alpha,p)q(\alpha,p)}$$

$$\frac{\partial q_0(\alpha,p)^2}{\partial p} = \frac{2\beta q_0(\alpha,p)q(\alpha,p)}{1+(\alpha-\mu+1)e^{\alpha-p\beta}}$$

Proof. Let α^* be the solution to $\sigma^2 = (\alpha-\mu)(1+e^{\alpha-p\beta})$ and note that $\alpha^* > \mu$ and let $q(p) = q(\alpha^*, p)$ and $q_0(p) = q_0(\alpha^*, p)$. Then, the approximate purchase probability is

$$\widetilde{s}(p) = \left[1+\sigma^2 q_0(p)q(p)\right]^{-1/2} q(p)e^{-\frac{(\alpha^*-\mu)^2}{2\sigma^2}}$$

and the corresponding revenue function $\widetilde{\pi}(p) = \widetilde{s}(p)p$, so that

$$\log\widetilde{\pi}(p) = \log p - \frac{1}{2}\log\left(1+\sigma^2 q_0(p)q(p)\right) + \log q(p) - \frac{(\alpha^*-\mu)^2}{2\sigma^2}$$

where we denote the four parts of this function by $f^{(I)}$, $f^{(II)}$, $f^{(III)}$, and $f^{(IV)}$, respectively. Consequently, we have that

$$\frac{\partial \log\widetilde{\pi}(p)}{\partial p} = \frac{\partial f^{(I)}}{\partial p} + \frac{\partial f^{(II)}}{\partial p} + \frac{\partial f^{(III)}}{\partial p} + \frac{\partial f^{(IV)}}{\partial p}$$

where

$$\frac{\partial f^{(I)}}{\partial p} = \frac{1}{p}$$

$$\begin{aligned}
\frac{\partial f^{(II)}}{\partial p} &= -\frac{\sigma^2}{2(1+\sigma^2 q_0(p)q(p))}\frac{\partial(q_0(p)q(p))}{\partial p} \\
&= -\frac{\sigma^2}{2(1+\sigma^2 q_0(p)q(p))}\left(\frac{\beta q_0(p)q(p)}{1+\sigma^2 q_0(p)q(p)} - 2q_0(p)\frac{\beta q_0(p)q(p)}{1+\sigma^2 q_0(p)q(p)}\right) \\
&= \frac{\sigma^2\beta q_0(p)q(p)(2q_0(p)-1)}{2(1+\sigma^2 q_0(p)q(p))^2}
\end{aligned}$$

$$\frac{\partial f^{(III)}}{\partial p} = \frac{1}{q(p)}\frac{\partial q(p)}{\partial p} = -\frac{\beta q_0(p)}{1+\sigma^2 q_0(p)q(p)}$$

$$\frac{\partial f^{(IV)}}{\partial p} = -\frac{\alpha^*-\mu}{\sigma^2}\frac{\partial \alpha^*}{\partial p} = -\frac{(\alpha^*-\mu)\beta q_0(p)q(p)}{1+\sigma^2 q_0(p)q(p)}$$

so that

$$\begin{aligned}
\frac{\partial \log \widetilde{\pi}(p)}{\partial p} &= \frac{1}{p} + \frac{\sigma^2\beta q_0(p)q(p)(2q_0(p)-1)}{2(1+\sigma^2 q_0(p)q(p))^2} - \frac{\beta q_0(p)}{1+\sigma^2 q_0(p)q(p)} - \frac{(\alpha^*-\mu)\beta q_0(p)q(p)}{1+\sigma^2 q_0(p)q(p)} \\
&= \frac{1}{p} + \frac{\beta q_0(p)}{1+\sigma^2 q_0(p)q(p)}\left[\frac{\sigma^2 q(p)(q_0(p)-1/2)}{1+\sigma^2 q_0(p)q(p)} - (\alpha^*-\mu)q(p) - 1\right] \\
&= \frac{1}{p} - \frac{\beta q_0(p)}{1+\sigma^2 q_0(p)q(p)}\left[\frac{\frac{1}{2}(\sigma^2 q(p)+2)}{1+\sigma^2 q_0(p)q(p)} + (\alpha^*-\mu)q(p)\right] \\
&= \frac{1}{p} - \frac{\beta q_0(p)}{1+\sigma^2 q_0(p)q(p)}\left[\frac{(\frac{1}{2}\sigma^2+\alpha^*-\mu)q(p)+\sigma^2(\alpha^*-\mu)q_0(p)q(p)^2+1}{1+\sigma^2 q_0(p)q(p)}\right] \\
&= \frac{1}{p} - \beta\frac{\overbrace{\left(\frac{\sigma^2}{2}+\alpha^*-\mu\right)q_0(p)q(p)+\sigma^2(\alpha^*-\mu)[q_0(p)q(p)]^2+q_0(p)}^{=f}}{\underbrace{[1+\sigma^2 q_0(p)q(p)]^2}_{=g}}.
\end{aligned}$$

To prove the claim, it suffices to show that this expression decreases in p. We will do so by showing that $\frac{\partial f}{\partial p}g - f\frac{\partial g}{\partial p} > 0$ under the hypothesis. Therefore, let

$$f^{(a)} = \left(\frac{\sigma^2}{2}+\alpha^*-\mu\right)q_0(p)q(p)$$

and

$$f^{(b)} = \sigma^2(\alpha^*-\mu)[q_0(p)q(p)]^2,$$

so that

$$\begin{aligned}
\frac{\partial f^{(a)}}{\partial p} &= \left(\frac{\sigma^2}{2}+\alpha^*-\mu\right)\frac{\partial(q_0(p)q(p))}{\partial p} + \frac{\partial\alpha^*}{\partial p}q_0(p)q(p) \\
&= \left(\frac{\sigma^2}{2}+\alpha^*-\mu\right)\frac{\beta q_0(p)q(p)(1-2q_0(p))}{1+\sigma^2 q_0(p)q(p)} + \frac{\beta\sigma^2 q_0(p)q(p)}{1+\sigma^2 q_0(p)q(p)}q_0(p)q(p) \\
&= \frac{\beta q_0(p)q(p)}{1+\sigma^2 q_0(p)q(p)}\left[\left(\frac{\sigma^2}{2}+\alpha^*-\mu\right)(1-2q_0(p)) + \sigma^2 q_0(p)q(p)\right]
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial f^{(b)}}{\partial p} &= \sigma^2\frac{\partial\alpha^*}{\partial p}[q_0(p)q(p)]^2 + 2\sigma^2(\alpha^*-\mu)q_0(p)q(p)\frac{\partial(q_0(p)q(p))}{\partial p} \\
&= \sigma^2[q_0(p)q(p)]^2\frac{\beta\sigma^2 q_0(p)q(p)}{1+\sigma^2 q_0(p)q(p)} + 2\sigma^2(\alpha^*-\mu)q_0(p)q(p)\frac{\beta q_0(p)q(p)(1-2q_0(p))}{1+\sigma^2 q_0(p)q(p)} \\
&= \frac{\beta\sigma^2[q_0(p)q(p)]^2}{1+\sigma^2 q_0(p)q(p)}\left[\sigma^2 q_0(p)q(p) + 2(\alpha^*-\mu)(1-2q_0(p))\right].
\end{aligned}$$

Consequently, we can obtain

$$\begin{aligned}\frac{\partial f}{\partial p} =& \frac{\partial f^{(a)}}{\partial p} + \frac{\partial f^{(b)}}{\partial p} + \frac{\beta q_0(p)q(p)}{1+\sigma^2 q_0(p)q(p)} \\ =& \frac{\beta q_0(p)q(p)}{1+\sigma^2 q_0(p)q(p)} \left[\left(\frac{\sigma^2}{2} + \alpha^* - \mu\right)(1-2q_0(p)) + \dots \right. \\ & \left. \sigma^2 q_0(p)q(p)\left(\sigma^2 q_0(p)q(p) + 2(\alpha^*-\mu)(1-2q_0(p)) + 1\right) + 1\right]\end{aligned}$$

and

$$\frac{\partial g}{\partial p} = 2\sigma^2 \beta q_0(p)q(p)(1-2q_0(p)).$$

We proceed by checking that $\frac{\partial f}{\partial p} g - f \frac{\partial g}{\partial p} > 0$, which is sufficient to prove the claim, i.e.,

$$\frac{\beta q_0(p)q(p)}{1+\sigma^2 q_0(p)q(p)} \left[\left(\frac{\sigma^2}{2} + \alpha^* - \mu\right)(1-2q_0(p)) + \sigma^2 q_0(p)q(p) + \right.$$
$$\left. \sigma^2 q_0(p)q(p)\left(\sigma^2 q_0(p)q(p) + 2(\alpha^*-\mu)(1-2q_0(p))\right) + 1\right][1+\sigma^2 q_0(p)q(p)]^2 -$$
$$\left[\left(\frac{\sigma^2}{2} + \alpha^* - \mu\right) q_0(p)q(p) + \sigma^2(\alpha^*-\mu)[q_0(p)q(p)]^2 + q_0(p)\right] 2\sigma^2\beta q_0(p)q(p)(1-2q_0(p)) > 0$$

$$\Leftrightarrow$$

$$(1+\sigma^2 q_0(p)q(p)) \left[\left(\frac{\sigma^2}{2} + \alpha^* - \mu\right)(1-2q_0(p)) + \right.$$
$$\left. \sigma^2 q_0(p)q(p)\left(1 + \sigma^2 q_0(p)q(p) + 2(\alpha^*-\mu)(1-2q_0(p))\right) + 1\right] -$$
$$2\sigma^2(1-2q_0(p)) \left[\left(\frac{\sigma^2}{2} + \alpha^* - \mu\right) q_0(p)q(p) + \sigma^2(\alpha^*-\mu)[q_0(p)q(p)]^2 + q_0(p)\right] > 0$$

$$\Leftrightarrow$$

$$(1-2q_0(p)) \left\{(1+\sigma^2 q_0(p)q(p)) \left(\frac{\sigma^2}{2} + \alpha^* - \mu + 2\sigma^2(\alpha^*-\mu)q_0(p)q(p)\right) - \right.$$
$$\left. 2\sigma^2 \left[\left(\frac{\sigma^2}{2} + \alpha^* - \mu\right) q_0(p)q(p) + \sigma^2(\alpha^*-\mu)[q_0(p)q(p)]^2 + q_0(p)\right]\right\} +$$
$$(1+\sigma^2 q_0(p)q(p))\left[\sigma^2 q_0(p)q(p)\left(1+\sigma^2 q_0(p)q(p)\right) + 1\right] > 0$$

$$\Leftrightarrow$$

$$\sigma^2(1-2q_0(p)) \left\{(1+\sigma^2 q_0(p)q(p)) \left(\frac{1}{2} + \frac{\alpha^*-\mu}{\sigma^2} + 2(\alpha^*-\mu)q_0(p)q(p)\right) - \right.$$
$$\left. 2\left(\frac{\sigma^2}{2} + \alpha^* - 2\mu\right) q_0(p)q(p) - 2\sigma^2(\alpha^*-\mu)[q_0(p)q(p)]^2 - 2q_0(p)\right\} +$$
$$(1+\sigma^2 q_0(p)q(p))\left[\sigma^2 q_0(p)q(p)\left(1+\sigma^2 q_0(p)q(p)\right) + 1\right] > 0$$

$$\Leftrightarrow$$

$$\sigma^2(1-2q_0(p)) \left\{\frac{1}{2} + \frac{\alpha^*-\mu}{\sigma^2} + \left(\alpha^* - \mu - \frac{\sigma^2}{2}\right) q_0(p)q(p) - 2q_0(p)\right\} +$$
$$(1+\sigma^2 q_0(p)q(p))\left[\sigma^2 q_0(p)q(p)\left(1+\sigma^2 q_0(p)q(p)\right) + 1\right] > 0 \Leftrightarrow$$
$$(\alpha^*-\mu)(1+e^{\alpha^*-p\beta})(1-2q_0(p)) \left\{\frac{1}{2} + \frac{1}{2}(\alpha^*-\mu)q_0(p)q(p)\left(1-e^{\alpha^*-p\beta}\right) - q_0(p)\right\} +$$
$$(\alpha^*-\mu)^3 q(p)^3 + 2(\alpha^*-\mu)^2 q(p)^2 + 2(\alpha^*-\mu)q(p) + 1 > 0$$

$$\Leftrightarrow$$

$$(\alpha^* - \mu)q_0(p)(1 - e^{\alpha^* - p\beta}) - \frac{1}{2}(\alpha^* - \mu)^2 q_0(p)q(p)\left(1 - e^{\alpha^* - p\beta}\right)^2 - \frac{1}{2}(\alpha^* - \mu)(1 - e^{\alpha^* - p\beta}) +$$
$$(\alpha^* - \mu)^3 q(p)^3 + 2(\alpha^* - \mu)^2 q(p)^2 + 2(\alpha^* - \mu)q(p) + 1 > 0$$
$$\Leftrightarrow$$
$$(\alpha^* - \mu)^3 q(p)^3 + (\alpha^* - \mu)^2\left(2q(p)^2 - \frac{1}{2}q_0(p)q(p)(1 - e^{\alpha^* - p\beta})^2\right) +$$
$$(\alpha^* - \mu)\left(q_0(p)(1 - e^{\alpha^* - p\beta}) - \frac{1}{2}(1 - e^{\alpha^* - p\beta}) + 2q(p)\right) + 1 > 0$$
$$\Leftrightarrow$$
$$(\alpha^* - \mu)^3 q(p)^3 + (\alpha^* - \mu)^2\left(\frac{7}{2}q(p)^2 - \frac{1}{2}q(p) - \frac{1}{2}q(p)^2 e^{\alpha^* - p\beta}\right) + \frac{1}{2}\sigma^2 + 1 > 0$$
$$\Leftrightarrow$$
$$2(\alpha^* - \mu)q(p)^3 + \left(7 - e^{\alpha^* - p\beta}\right)q(p)^2 - q(p) + \frac{2 + (\alpha^* - \mu)(1 + e^{\alpha^* - p\beta})}{(\alpha^* - \mu)^2} > 0$$
$$\Leftrightarrow$$
$$2(\alpha^* - \mu)q(p) + \frac{2 + (\alpha^* - \mu) + (\alpha^* - \mu)e^{\alpha^* - p\beta}}{q(p)^2(\alpha^* - \mu)^2} + \frac{q_0(p)}{q(p)} - \frac{q(p)}{q_0(p)} + 6 > 0$$
$$\Leftrightarrow$$
$$2\sigma^2 q_0(p)q(p) + \frac{2 + \sigma^2}{\sigma^4 q(p)^2 q_0(p)^2} + \frac{q_0(p)}{q(p)} - \frac{q(p)}{q_0(p)} + 6 > 0$$
$$\Leftrightarrow$$
$$2\sigma^2 q_0(p)^3 q(p)^3 + 6q_0(p)^2 q(p)^2 + q_0(p)^3 q(p) - q_0(p)q(p)^3 + \frac{1}{\sigma^2} + \frac{2}{\sigma^4} > 0$$

which holds for $\sigma^2 < 1$. Now suppose $\sigma^2 \geq 1$, then

$$2\sigma^2(q(p) - q(p)^2)^3 + 6q(p)^4 - 10q(p)^3 + 3q(p)^2 + q(p) + \frac{1}{\sigma^2} + \frac{2}{\sigma^4}$$
$$>$$
$$q(p) + 3q(p)^2 - 8q(p)^3 + 6q(p)^5 - 2q(p)^6 + \frac{1}{\sigma^2} + \frac{2}{\sigma^4}.$$

Now, by using Sturm's theorem one can show that for $\sigma^2 \leq 34.92$ the last expression has no roots for $q(p) \in [0, 1]$. Together with the observation that for any values $q(p) \in [0, 1]$ and $\sigma^2 \in [0, 34.92]$ this expression is positive, leads to

$$q(p) + 3q(p)^2 - 8q(p)^3 + 6q(p)^5 - 2q(p)^6 + \frac{1}{\sigma^2} + \frac{2}{\sigma^4} > 0$$

for $\sigma^2 \leq 34.92$. □

References

Akcay, Y., Natarajan, H. P., and Xu, S. H. (2010). Joint dynamic pricing of multiple perishable products under consumer choice. *Management Science*, 56(8):1345–1361.

Alptekinoğlu, A. and Semple, J. H. (2016). The exponomial choice model: A new alternative for assortment and price optimization. *Operations Research*.

Asmussen, S., Jensen, J. L., and Rojas-Nandayapa, L. (2014). On the Laplace transform of the lognormal distribution. *Methodology and Computing in Applied Probability*, pages 1–18.

Aydin, G. and Porteus, E. L. (2008). Joint inventory and pricing decisions for an assortment. *Operations Research*, 56(5):1247–1255.

Butler, R. W., Wood, A. T., et al. (2002). Laplace approximations for hypergeometric functions with matrix argument. *The Annals of Statistics*, 30(4):1155–1177.

Dong, L., Kouvelis, P., and Tian, Z. (2009). Dynamic pricing and inventory control of substitute products. *Manufacturing & Service Operations Management*, 11(2):317–339.

Gallego, G. and Wang, R. (2014). Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Operations Research*, 62(2):450–461.

Hanson, W. and Martin, K. (1996). Optimizing multinomial logit profit functions. *Management Science*, 42(7):992–1003.

Harding, M. C. and Hausman, J. (2007). Using a Laplace approximation to estimate the random coefficients logit model by nonlinear least squares. *International Economic Review*, 48(4):1311–1328.

Hensher, D. A. and Greene, W. H. (2003). The mixed logit model: the state of practice. *Transportation*, 30(2):133–176.

Huh, W. T. and Li, H. (2015). Technical note — pricing under the nested attraction model with a multistage choice structure. *Operations Research*, 63(4):840–850.

Li, G., Rusmevichientong, P., and Topaloglu, H. (2015). The d-level nested logit model: Assortment and price optimization problems. *Operations Research*, 63(2):325–342.

Li, H. and Huh, W. T. (2011). Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing & Service Operations Management*, 13(4):549–563.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, pages 105–142.

McFadden, D. (1980). Econometric models for probabilistic choice among products. *Journal of Business*, pages S13–S29.

McFadden, D. et al. (1978). *Modelling the choice of residential location.* Institute of Transportation Studies, University of California California.

McFadden, D., Train, K., et al. (2000). Mixed MNL models for discrete response. *Journal of applied Econometrics*, 15(5):447–470.

Rossi, P. E., McCulloch, R. E., and Allenby, G. M. (1996). The value of purchase history data in target marketing. *Marketing Science*, 15(4):321–340.

Rusmevichientong, P., Shmoys, D., Tong, C., and Topaloglu, H. (2014). Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management*, 23(11):2023–2039.

Song, J.-S. and Xue, Z. (2007). Demand management and inventory control for substitutable products. *Working paper*.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.